

Patent Application
for
OBJECT RECOGNITION SYSTEM

5 Field of the Invention

The present invention relates to an optical object recognition system which detects objects in front of a vehicle such as an automobile, etc., using one or more sensors mounted on this vehicle. More specifically, the present invention
10 concerns an object recognition system which recognizes objects using a plurality of windows in the captured images.

Background of the Invention

15 In recent years, devices which determine the distance and the size of objects in front of a vehicle, and which appropriately control the vehicle in accordance with this determination, have been proposed for the purpose of improving the safety of vehicle operation.

20 Japanese Patent Application Kokai No. Hei 7-225126 describes one example of a device for recognizing objects in front of a car appropriately. The device utilizes an optical distance measuring device consisting of two light-receiving elements, and determines whether an object whose distance has been detected is a physical object (an obstacle) or a road area (including characters or
25 white lines on the road surface). Specifically, the device comprises a stereo camera capturing an image of an object on the road, divides the image into a plurality of windows, and calculates the distance between the stereo camera and the captured object for each window. The object in front of a car is recognized

by comparing the distance with a reference distance determined for each row range of the windows.

The device disclosed in Japanese Patent Application Kokai No. Hei 7-225126 decides whether the object is a physical object or a road area based on an assumption that the road spreads horizontally ahead of the vehicle. Difference may exist between the actual road and the road estimated based on the assumption when factors such as an attachment position of the camera or an installation angle have deviated from the predetermined specification.

This makes it difficult to judge whether a measured value truly represents the distance to the road and may lead to a false recognition of the object. Specifically, difference may occur in such occasions as when the vehicle is lopsided by pitching or rolling, when a vehicle travels along a hill, or when the attachment point or the installation angle of the camera has changes from the location as initially specified.

Summary of the Invention

Accordingly, one object of the present invention is to provide a system, which recognizes objects in front of a vehicle even when the vehicle is lopsided by pitching or rolling or when it is traveling up or down hill, or the camera has moved from a specified point or a specified angle.

In according with one aspect of the invention, an object recognition system, mounted on a vehicle comprises an image sensor and a controller. The sensor captures an image of an object. The controller divides the image into a plurality of windows. The controller measures a distance to the road surface for each of the plurality of windows and thereby obtains a plurality of distances. The

controller estimates a relative inclination of the road surface against the vehicle based on the plurality of distances. The controller judges, for each of the plurality of windows, based on the estimated inclination, whether the object is an obstacle or the road surface and thereby obtains a judgment result. The controller
5 recognizes the object based on the judgment result.

In accordance with another aspect of the invention, an object recognition method for recognizing an object ahead of a vehicle comprises the steps of capturing an image in front of the vehicle, measuring a distance to the road
10 surface for each of a plurality of windows dividing the image, estimating a relative inclination of the road surface against the vehicle based on the plurality of distances, judging based on the estimated inclination whether the object is an obstacle or the road surface for each of the windows, and recognizing the object based on such judgment.

15 In accordance with further another aspect of the invention, an object recognition system comprises one or more sensors for capturing an image of an object, measuring means for measuring a distance to the road surface for each of a plurality of windows dividing the image and thereby obtains a plurality of
20 distances, estimation means for estimating a relative inclination of the road surface against the vehicle based on the plurality of distances, judging means for judging based on the estimated installation angle whether the object is an obstacle or the road surface for each of the plurality of windows, and recognition means for recognizing the object based on a result from the judging means.

25 According to an aspect of the invention, an estimation of a relative inclination between a vehicle and a road surface is possible even when the vehicle is lopsided by pitching (tilt in a front-to-rear direction due to acceleration

or deceleration) or rolling (tilt in a side-to-side direction at a curve), or when the vehicle travels along a hill or a road with banks. The estimated inclination allows judgment whether the object is an obstacle, or a character or a sign on the road, thus enabling an appropriate recognition of the object.

5

According to another aspect of the invention, an object recognition system mounted on a vehicle comprises one or more image sensors for capturing an image of an object, measuring means for measuring distance to a road surface for each of a plurality of windows dividing the image, estimation means for
10 estimating a relative inclination of the road surface against the vehicle based on the measured distances, correction means for correcting position of the image sensor based on the estimated inclination, judging means for judging based on the corrected position of the image sensor whether the object is an obstacle or a road surface for each of the windows, and recognition means for recognizing the
15 object based on judgment result from the judging means.

In accordance with such an aspect of the invention, an object can be properly recognized even when positioning error of the image sensor has accumulated due to secular changes over a long period of time. The positioning
20 error of the image sensor, typically a camera, may be estimated and corrected based on measurement of the distance to the road surface.

Brief Description of the Drawings

25 Figure 1 is a block diagram illustrating an overall structure of one embodiment of the present invention.

Figure 2 is a diagram illustrating the principle of measurement by a triangulation method.

Figure 3(a) is a diagram showing an image that is captured, and Figure 3(b) shows the image divided into small areas (windows) for the purpose of judging distances and road areas, in accordance with one embodiment of the present invention.

5 Figure 4(a) and Figure 4(b) are diagrams in the case of pitching. Figure 4(a) is showing an actual road, an estimated road and the deleted ranges of the road, and Figure 4(b) shows the content of the three-dimensional road memory plotted on the z-y plane.

Figure 5(a) and Figure 5(b) are diagrams in the case of rolling.

10 Figure 5(a) is showing an actual road, an estimated road and deleted ranges of the road, and Figure 5(b) shows the content of the three-dimensional road memory plotted on the x-y plane.

Figure 6 is a diagram illustrating a road estimated from the content of the road pitch memory and the road roll memory.

15 Figure 7 is a diagram illustrating the coordinate systems related to the equation indicating the correlation between the position on the CCD plane and the distance to the road.

Detailed Description of the Invention

20 The invention will now be described relative to a preferred embodiment referring to the attached drawings. Figure 1 is an overall block diagram of an object recognition system in accordance with one embodiment of the present invention. Figure 2 is a diagram which indicates the distance measurement principle based on the triangulation method used in the present embodiment.
25 First, a distance measurement method using a pair of image sensors will be described with reference to Figure 2.

A line sensor 21 and lens 23 constituting one of a pair of image sensors in a stereo camera are installed at a specified distance, i. e., at a distance equal to the base line length B in the horizontal or vertical direction from the line sensor 22 and lens 24 constituting the other of a pair of image sensors. The line sensors 21 and 22 are typically one-dimensional CCDs, but may also be linearly arranged photo-sensor arrays. For the use at night, image sensors using infrared light are advisable. In this case, it is advisable to install infrared-transparent filters in front of the lenses 23 and 24, and to devise the system so that an object 20 is illuminated at predetermined time intervals using an infrared light source.

10 Infrared light reflected from the object 20 is sensed by the line sensors 21 and 22.

The line sensors 21 and 22 are respectively positioned at the focal length "f" of the lenses 23 and 24. Assuming that an image of an object located at distance "a" from the plane of the lenses 23 and 24 is formed at a position shifted by a distance X1 from the optical axis of the lens 23 in the case of the line sensor 21, and is formed at a position shifted by a distance X2 from the optical axis of the lens 24 in the case of the line sensor 22, then, according to the principle of triangulation, the distance "a" to the object 20 from the plane of the lenses 23 and 24 is determined by the equation: $a = B \cdot f / (X1 + X2)$.

20 In the present embodiment, the images are digitized. And accordingly, the distance (X1 + X2) is digitally calculated. The sum of the absolute values of the differences between the digital values indicating the brightness of the corresponding pixels of both images obtained from the line sensors 21 and 22 is determined while one or both of the images are shifted, and this sum is taken as a correlation value. The amount of shift of the images when this correlation value is at a minimum indicates the positional deviation between the two images, i. e., (X1 + X2). In idealized terms, the distance by which the two images obtained

from the line sensors 21 and 22 must be moved in order to cause the images to overlap as shown in Figure 2 is $(X1 + X2)$.

Here, for the sake of simplicity, the image sensors were described as one-dimensional line sensors 21 and 22. However, in one embodiment of the present invention, as will be described below, two-dimensional CCDs or two-dimensional photo-sensor arrays are used as image sensors. In this case, the same correlation calculations as those described above are performed by relatively shifting the two-dimensional images obtained from the two image sensors. The amount of shift at the point where the correlation value reaches a minimum corresponds to $(X1 + X2)$.

The image sensor 3 shown in Figure 1 corresponds to one of the image sensors in Figure 2, comprising the lens 23 and line sensor 21, and the image sensor 3' corresponds to the other image sensor in Figure 2, comprising the lens 24 and line sensor 22. In this embodiment, as is shown in Figure 3 (b), the imaged area is divided into a plurality of windows (small sections) W_{11}, W_{12}, \dots and distance is measured for each window. Accordingly, a two-dimensional image of the overall object is required. Accordingly, each of the image sensors 3 and 3' is comprised of a two-dimensional CCD array or a two-dimensional photo-sensor array.

Figure 3 (a) shows an example of the image obtained when another vehicle running in front of the vehicle incorporating the system of the present invention is imaged by one of the image sensors 3 and 3'. Figure 3 (b) shows the image in Figure 3 (a) schematically split into a plurality of small sections called windows. Figure 3 (b) has rows in the vertical direction and columns in horizontal direction. For the sake of simplicity, the image is shown split into 10

rows \times 15 columns of windows. Reference numerals are assigned to the respective windows. For example, W_{12} indicates the window in row 1, column 2.

Referring to Figure 1, the images of objects captured by the image sensors 3 and 3' are converted into digital data by analog-digital converters (A/D converters) 4 and 4' and stored in image memories 5 and 5'. The image portions corresponding to the window W_{11} are respectively cut out from the image memories 5 and 5' by a window cutout part 13 and sent to a correlation calculating part 6. The correlation calculating part 6 shifts the two cutout images by a specified unit at a time, and performs the aforementioned correlation calculations. The amount of shift at the point where the correlation value reaches a minimum corresponds to $(X1 + X2)$. The correlation calculating part 6 sends the value of $(X1 + X2)$ thus determined to a distance (DT) calculating part 7.

The distance calculating part 7 determines the distance a_{11} to the object in the window W_{11} using the aforementioned formula: $a = B \cdot f / (X1 + X2)$. The distance a_{11} thus determined is stored in a distance memory 8. A similar calculation process is successively performed for respective windows, and the resulting distances a_{11}, a_{12}, \dots are stored in the distance memory 8. The distance to a captured object calculated for each window is referred to as the "measured distance" of the window.

In the image data used in the above-mentioned correlation calculations, the pitch of the elements in the imaging element array determines the resolution. Accordingly, when a light-receiving element such as a photo-sensor array that has a relatively large pitch is used, it is preferred to enhance the density of the image data by performing calculations involving inter-pitch interpolation.

5dp
5/18/03

5

20

25

without being lopsided. Such values may be calculated from the attachment position of the CCD cameras, the installation angle, the base line length B, the focal distance f, the size of the CCD and the positions of the windows in the image. The estimated distance memory 32 stores the initial values for respective
5 windows.

In the above example, the measured distance values of windows that have been determined to belong to the road surface are deleted from the distance memory 8. Alternatively, flags may be set for such windows identifying that the
10 windows belong to the road surface. Such windows with flags may be stored in the distance memory 8 thus omitting the road surface distance memory 51.

A three-dimension converter 52 converts the positions of windows and the corresponding measured distance values into coordinates in a three-dimensional
15 space, with the origin point being the attachment position of the camera, the y-axis being vertical downward direction, the z-axis being the direction of the vehicle's straight travel, and the x-axis being perpendicular to both the y and z axes. The x-axis represents the width direction of the vehicle. The converted coordinates are stored in a three-dimensional road surface memory 53.

20 An object recognition system utilizing three-dimensional conversion of the image is described in U.S. patent application Serial No. 09/572,249 filed on May 17, 2000, which is incorporated herein by reference.

25 A road surface pitch estimator 9 estimates relative inclination between the camera and the road surface from the three-dimensional positions stored in the road surface memory 53 in the case of tilt in the front-to-rear direction due to

pitching (drop in the rear or the front of the vehicle in acceleration or deceleration) or a travel along a hill.

5 Figs. 4(A) and 4(B) show the correlation between an estimated road surface S1 and an actual road surface S2 in the case of front-dropping pitching. The actual road surface S2 represents the relative road surface on the assumption that the vehicle is positioned horizontally. In that case, the relative road surface appears as an uphill slope.

10 In Fig. 4(A), a line S1 represents a virtual road surface that is positioned above the actual road surface by a predetermined distance such as 10cm. The window whose measured distance value are below the surface S1 is regarded as capturing the road surface and its position and the distance value are stored in the road surface distance memory 51 by the road surface exclusion part 31.

15 Those position and distance are converted through the three-dimensional converter 52 into a three-dimensional position and stored in the three-dimension road surface memory 53.

20 Fig. 4(B) shows the three-dimensional position stored in the memory 53, projected onto the z-y plane. The road surface pitch estimator 9 samples those positions in the memory 53 whose respective z-coordinate values are equal to or lower than a predetermined value, for example, 15m. If the number of the samples is equal to or larger than a predetermined value, for example, 200, equations (1) to be described hereafter is used to determine a straight line

25 representing best the correlation of the samples projected onto the z-y plane. The inclination, the y-intercept and the number of the samples of the straight line are stored in a road surface pitch memory 54. On the other hand, the number of the samples less than 200 leads to reservation of the content of the memory 54

without its content being updated. The inclination of the straight line represents the inclination of the road surface in the front-rear direction in Fig. 4(A) and the y-axis intercept represents the height of the camera from the road surface.

$$\tan \theta = \frac{n \sum Z_i Y_i - \sum Z_i \sum Y_i}{n \sum Z_i^2 - (\sum Z_i)^2} \quad (1)$$

$$h_p = \frac{\sum Y_i - \tan \theta \sum Z_i}{n}$$

“Y_i” and “Z_i” respectively represent the y and z coordinates of the i-th sample position, “n” the number of the samples, “tan θ” the slope of the straight line and “h_p” the y-axis intercept.

A road surface roll estimator 10 estimates a relative inclination of the road surface compared with the camera. The estimation is performed utilizing the three-dimensional position stored in the memory 53 in the case of tilt of the vehicle in the side-to-side direction due to rolling (drop in left-side or right-side of the vehicle in the motion of cornering), running a road with a bank thereon, or else.

Figs. 5(A) and 5(B) show the correlation between an estimated road surface S1' and an actual road surface S2' in the case of right-side-drop rolling. The actual road surface S2' represents the relative road surface on the assumption that the vehicle is horizontally positioned. In that case, the relative road surface appears as a right-side-up slope.

In Fig. 5(A), a dotted line S1' represents a virtual road surface that is positioned above the actual road surface by a predetermined distance, for

example, 10cm. The window whose measured distance value is lower than the imaginary surface S1' is regarded as capturing the road surface, and its position and the distance value are stored in the road surface distance memory 51 by the road surface exclusion part 31. Those position and distance are converted
 5 through the three-dimensional converter 52 into a three-dimensional position and stored in the three-dimension road surface memory 53.

Fig. 5(B) shows the three-dimensional position stored in the memory 53, projected onto the x-y plane. The road surface roll estimator 10 samples those
 10 positions in the memory 53 whose respective z-coordinate values are equal to or lower than a predetermined value, for example, 15m. When the number of the samples is equal to or more than a predetermined value, for example, 200, equations (2) to be described hereafter is used to determine a straight line best representing the correlation of the samples projected onto the x-y plane. The
 15 slope, the intercept of the y-axis and the number of the samples of the straight line are stored in a road surface roll memory 55. On the other hand, when the number of the samples is less than 200, the content of the memory 55 is reserved without being updated. The slope of the straight line represents the slope in the side-to-side direction of Fig. 4(A), and the y-axis intercept represents the height
 20 of the camera from the road surface.

$$\tan \alpha = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} \quad (2)$$

$$h_r = \frac{\sum Y_i - \tan \alpha \sum X_i}{n}$$

25 “Xi” and “Yi” respectively represent the x and y coordinates of the i-th sample position, “n” the number of the samples, “tan α” the slope of the straight line and “hr” the y-axis intercept.

The estimators 9, 10 sample only those positions whose respective z-coordinates are equal to or less than 15m because the further the distance is, the lower the accuracy of the positions is, and the further the distance is, the larger the discrepancies between the actual road surface and the estimated road surface will grow, which may lead to capturing of other objects than the road surface, which in turn would cause errors in the estimation if added to the samples. Also, the reason for not carrying out slope estimation when the number of the samples is less than 200 is that smaller samples makes estimated values not credible.

An estimated distance corrector 11 corrects stored content of the estimated distance memory 32 according to the respective slopes in the front-to-rear direction and the side-to-side direction and the height of the camera stored in the memory 54 and 55, so that the content represents more appropriate distances to the road surface.

The correlation between the position (x, y) on the plane of the CCD to capture the image of the road surface and the distance to the road surface may be estimated utilizing equations (3) even in the case of pitching, rolling or running a hill slope or a road with banks.

$$h = (h_p + h_r) / 2$$

$$\varepsilon = x \tan \alpha - y - F \tan \theta \quad (3)$$

$$d = h \sqrt{y^2 + F^2} / \varepsilon \quad (\varepsilon > 0),$$

$$0 \quad (\varepsilon \leq 0)$$

5 "tan θ " and "tan α " respectively represent the slopes of the straight lines
 respectively stored in the memory 54 and 55, and "hp" and "hr" the y-axis
 intercepts respectively in the memory 54 and 55. "F" represents the focal
 distance of the CCDs. "d = 0" implies that windows are not capturing the road
 surface.

10 As shown in Fig. 6, the slope and intercept of the line stored in road
 surface pitch memory 54 and the slope and intercept of the line stored in the road
 surface roll memory 55 are combined to estimate a single plane. The coordinate
 system of the CCD is shown in Fig. 7. The correlation between a position in the
 CCD plane in Fig. 7 onto which the plane in Fig. 6 is projected and a distance
 from the CCD to this plane in Fig. 6 are as shown in the equations (3).

15 The estimated distance corrector 11 estimates distances "d"s for respective
 windows according to the equations (3) to replace previously estimated distances
 already stored in the memory 32 with the newly estimated distances "d"s.

20 In the abovementioned example, the content of the memory 32 is replaced
 directly with distances "d"s, but another scheme may be adopted. For instance,
 weighed averages between the respective previous distances and the respective
 current distances are obtained to reduce the variation in estimation due to the
 measurement error. Moreover, as a further alternative, some sets of previously
 estimated distances are reserved over some period of time, and the current slope
 of the road and the current height of the camera may be estimated from the time
 25 sequence information.

In accordance with another aspect of the invention, use of a large time
 constant with which the content of the memory 32 is changed enables the system

to learn a secular change in the installation angle or attachment position error of the camera and to correct such change or error. This is due to the fact that, over a long period of time of travel by a vehicle, an average inclination of the road surface is substantially horizontal. Thus, average inclination due to pitch and roll of the vehicle is nearly zero in a long period of time or a large mileage. The average inclination in a long period of time or a large mileage indicates positioning error of the image sensor caused by secular change of physical alignment. Thus, an actual attachment position and an actual angle of the camera can be estimated based on the inclination angles and the heights stored in the memory 54 and 55.

A large time constant here means that the content of the road surface distance memory 32 is updated at an interval of three or six months, or is updated at an interval in terms of mileage, say every 1000 or 2000 miles.

A camera installation angle estimator 61 and a camera installation angle memory 62 may be separately provided. The camera installation angle estimator 61 calculates moving average of the contents in the memory 54 and 55, which have been accumulated over a long period of time, to estimate the average inclination of the road surface and the average height from the road surface to the camera and to store them in the memory 62.

In the above-mentioned example, the number of the samples less than 200 leads to the reservation of the contents in the memory 54 and 55. Instead, the contents of the memory 32 may be varied to approach its initial state. Or, in the case the estimator 61 and the memory 62 are provided, the content in the memory 32 may be operated to approach or reach a distance value estimated based on the height and the angle in the memory 62, whereby discrepancy is

avoided between the content in the memory 32 and the actual state of the road surface.

As described above, the estimated distance corrector 11 corrects the
5 estimated distances in the estimated distance memory 32 according to the
contents of the road surface pitch memory 54 and the road surface roll memory
55. This correction may be performed, in a hardware structure, in response to a
signal from estimator 9 or estimator 10 indicating that the level of change has
10 took place in the content of the pitch memory 54 or the roll memory 55 that
requires correction of the estimated distance stored in the memory 32. Or, in a
software structure, a central processing unit (CPU) controlling the object
recognition system may execute a correction routine at predetermined intervals,
for example, 67ms to correct the estimated distances.

15 The road surface exclusion part 31 compares the measured distance of
each window with corresponding corrected estimated distance in the memory 32.
Then a window with a measured distance larger than the corresponding corrected
distance minus a predetermined threshold value (e.g. 10cm) is considered to
have captured the road surface. On the other hand, the window with a measured
20 distance equal to or less than the corresponding corrected distance minus the
threshold value is considered to have captured a physical object other than the
road surface.

Thus, the road surface exclusion part 31 transfers the data of windows that
25 have captured physical objects other than the road surface to the object (OB)
recognition part 38. The object recognition part 38 recognizes one or more
objects based on the transferred data. The CPU performs the process for object
recognition at a predetermined interval, for example, 67ms. A consolidated

recognition based on the results obtained from 2 to 3 cycles of the recognition process will improve recognition accuracy.

An object memory 39 stores information about distance, position, size and the like of the recognized object. The object recognition part 38 calculates the relative speed between the object and the vehicle based on the two pieces of information respectively obtained from the previous and the current recognitions of the same object and then stores the speed in the object memory 39 as a part of current data related to the recognized object.

Reading the object information from the memory 39, a vehicle (VC) controller 45 accordingly executes a collision-prevention action such as sounding an alarm to remind the driver being too close to a vehicle in the front, or executing compulsive braking procedures by sending a signal to an electronic control unit for the engine or a brake control unit. At that time, the vehicle controller 45 receives the vehicle's own speed from an own-speed detector 46 and a yaw rate signal from a yaw rate detector 47 to decide the area to travel, and to keep a safe distance from the object by controlling the vehicle.

The correlation calculating part 6, the distance calculating part 7, distance memory 8, the window cutout part 13, the road surface exclusion part 31, the estimated distance memory 32, the object recognition part 38, the object memory 39, the image memory 5, 5', the road surface distance memory 51, the three-dimension converter 52, the three-dimension road surface memory 53, the road surface pitch estimator 9, the road surface roll estimator 10, the road surface pitch memory 54, the road surface roll memory 55, the camera installation angle estimator 61, the camera installation angle memory 62 and the vehicle controller 45 may be composed of a central processing unit (CPU), a read-only memory

storing control programs and control data, and a random access memory (RAM) providing a working area for the CPU and storing various data temporarily. The divided memory areas of a single RAM may be assigned to realize the distance memory 8, the estimated distance memory 32 and the object memory 39. Also, a
5 temporary memory area required for data storage in various operations may be provided utilizing a part of the RAM area.

Moreover, an output from the object recognition system of the invention LAN connected to an electric control unit (ECU) for an engine, an ECU for
10 brake control and additional ECU would be utilized for overall control of the vehicle.

According to the recognition system, as can be seen from the description, correct recognition is performed even when a relative road surface inclination
15 changes due to pitching or rolling, or travel on a hill or a road with a bank. Moreover, positioning error of the sensors due to secular change can be corrected.